DOCUMENT RESUME

ED 449 268 UD 033 955

AUTHOR	Howell, William G.; Wolf, Patrick J.; Peterson, Paul E.; Campbell, David E.
TITLE	The Effect of School Vouchers on Student Achievement: A Response to Critics. Occasional Paper.
INSTITUTION	Harvard Univ., Cambridge, MA. Kennedy School of Government.
PUB DATE	2000-00-00
NOTE	8p.; For the original study, see UD 033 744. Produced by the Program on Education Policy and Governance, located within the Taubman Center on State and Local Government, John F. Kennedy School of Government and within the Center for American Political Studies, Department of Government, Harvard University.
PUB TYPE	Opinion Papers (120) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; Educational Improvement; *Educational Vouchers; Elementary Secondary Education; High Risk Students; Hispanic Americans; *Minority Group Children; Private Education; Program Effectiveness; Public Education; Scores
IDENTIFIERS	African Americans

ABSTRACT

This paper responds to critics of an August 2000 paper that reported the effects of educational vouchers on student test scores in three cities. The report found that: (1) after 2 years, no students other than African Americans seemed to benefit from vouchers; and (2) African Americans in all three cities posted moderately large test score gains after 2 years. This paper responds to such criticisms as: the experimental group may have been biased, since some of the most disadvantaged voucher winners did not switch to private schools and were excluded, and the key finding improperly compared two dramatically different groups and may reflect private school screening-out of the most at-risk students; purported gains for African Americans are overstated; gains displayed by African American children are most distinct during their first year in a private school, but then the achievement advantage, relative to their peers in public schools, levels off; the demonstrated gains from using vouchers are limited to math, as reading performance appears to be more difficult to budge; and reported test score gains may be due to the declining share of students who appeared for the standardized tests from years 1 and 2. These criticisms are based upon an inaccurate characterization of the analysis. They misunderstand the design of the study and incorrectly suggest that it drops some students from the analysis. (SM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 449 268

The Effect of School Vouchers on Student Achievement:

A Response to Critics

By

William G. Howell, University of Wisconsin-Madison
Patrick J. Wolf, Georgetown University
Paul E. Peterson, Harvard University
David E. Campbell, Harvard University

Occasional Paper, Program on Education Policy and Governance, Harvard University

The Program on Education Policy and Governance (PEPG) is located within the Taubman Center on State and Local Government, John F. Kennedy School of Government and within the Center for American Political Studies, Department of Government, Harvard University. Mailing address: Taubman 306, Kennedy School of Government, 79 J. F. Kennedy St., Harvard University, Cambridge, MA 02138 Phone: 617-495-7976/495-8312; Fax: 617-496-4428; Website: http://data.fas.harvard.edu/pepg/

UD033955

BEST COPY AVAILABLE

A M Wendland
Prog. on Education Policy & Governance, Harvard Univ.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The Effect of School Vouchers on Student Achievement: A Response to Critics

On August 28, 2000, we released a paper to be presented before the annual meetings of the American Political Science Association, which reported on the effects of vouchers on student test scores in three cities-New York City, Dayton, Ohio, and Washington, D.C. The report presented two basic findings. First, after two years, no students other than African Americans appeared to benefit from the voucher programs. Latinos in New York and whites in Dayton who switched from a public to a private school did not score significantly higher or lower than their public school peers. In D.C., the comparison is inappropriate given the small number of non-African Americans in the study.

The second finding, which has attracted considerably more attention, is that African Americans in all three cities posted moderately large test score gains after two years. In New York, African Americans who switched from public to private schools scored 4 percentile points higher than the control group in their combined reading and math scores. In Dayton and D. C., they scored 6 and 9 percentile points higher, respectively. The results in all three cities are statistically significant.

In all three cities, the evaluations were designed as randomized field trials, what Caroline Hoxby of Harvard University calls the "gold standard" of social science research. Nonetheless, since releasing our report, a number of interest groups and scholars have leveled criticisms. Some of these criticisms raise important scholarly issues and deserve a response. Consider the following:

"The experimental group may have been biased as some of the most disadvantaged voucher winners did not switch to a private school, and therefore were excluded from the group (possibly boosting mean achievement levels artificially)" *PACE*, p. 10.

"The Peterson study's key finding improperly compares two dramatically different groups and may well reflect private school screening out of the most at-risk students" *People for the American Way*, p. 3.

These criticisms are based upon an inaccurate characterization of our analysis. They misunderstand the design of the study and incorrectly suggest that we drop some students from the analysis.

In the three cities, roughly half the students took the voucher that was offered to them (the takers) and about half did not (the decliners). As we state clearly in our reports, takers and decliners differed in a number of respects. Most notably, takers had higher family incomes in New York and D.C., but lower incomes in Dayton. The New York and D.C. findings are not surprising, given that the voucher awards did not cover all the costs of a private education. These additional costs were the reason most frequently given by families for not using the voucher. Presumably, take-up rates would rise if the monetary value of vouchers were increased.

We do not, however, drop the decliners from the analysis. All members of the control and treatment groups were invited to follow-up testing sessions, and every one of these families who showed up is included in the analysis. To estimate the impact of switching from a public to a private school, we do not simply compare takers and members of the control group, as the *PACE* report contends. Indeed, the very reason for presenting tables comparing takers and decliners in our reports is to justify the need for statistical models that account for any systematic differences between these groups.

In the absence of randomized field trials, analysts usually attempt to address this problem by controlling for initial test scores and family background characteristics. Such studies, however, are often criticized for not adequately controlling for unobserved differences between treatment and control groups.

Given the design of the voucher programs in New York, Dayton, and D.C., our research avoids many of the statistical problems associated with analyzing observational data. The sophisticated, widely used instrumental-variable model that we employ effectively adjusts for differences between takers and decliners. This analytical technique takes advantage of the fact that vouchers were offered at random, and thereby eliminates the bias introduced by differential take-up rates. The technique was first used in medical research, is now commonplace in econometric studies, and was employed by Alan Krueger in his analysis of the effects of class size on student performance in Tennessee, a study praised by many of the same people who have criticized our report.

Purported gains for African Americans are "overstated." Kate Zernike, *New York Times*.

To substantiate this claim, Zernike relies almost entirely upon a separate press release issued by Mathematica Policy Research (MPR) on the New York evaluation.

David Myers, senior fellow at MPR and co-principal investigator of the New York evaluation, expressed concern that results in New York City are not sufficiently consistent across grade levels to warrant the conclusion that voucher impacts have been detected. He agrees that statistically significant, positive effects of 4 percentile points on the test scores of all African Americans were observed in the New York evaluation. He also endorses the statistical apparatuses we use to evaluate the impacts of vouchers on student achievement. Myers points out, however, that when examined by grade level, statistically significant effects are limited to sixth graders in New York and, as a result, he concludes that there was "no impact." Myers was not involved in either the D.C. or Dayton evaluations.

It is worth highlighting that these fluctuations are limited to New York. Impacts found in Dayton and D.C. are not concentrated in any particular grade level. The test scores of middle-school African Americans who switched to a private school in D.C. did drop after the program's first

year; one year later, however, all African Americans in D.C. who switched to a private school, young and old alike, posted positive gains. In Dayton, the positive impacts observed in both years held for African Americans in multiple grade levels. When considering all three cities, the preponderance of the evidence suggests that after two years, African Americans as a group appeared to benefit from switching to a private school, while members of other ethnic groups did not.

At least after two years, then, the concentration of gains in a particular grade in New York appears exceptional. The finding, though, is hardly surprising. Random fluctuations often occur when one breaks down a sample and examines data grade by grade. For this reason, the education statistician Anthony Bryk, together with his colleagues, recommend that conclusions about school impacts not be drawn from "only single grade information... Judging a school by looking at only selected grades can be misleading. We would be better off, from a statistical perspective, to average across adjacent grades to develop a more stable estimate of school productivity."

Bryk et al's admonition is particularly compelling when, as is the case in New York, only 50 to 75 African American students are observed in the treatment and control groups at each grade level after two years. Under these circumstances, separate analyses run on individual grade levels are unlikely to generate stable estimates. Rather than focusing exclusively on inconsistencies between grade-specific findings in New York, we would do better to survey the full range of evidence collected from all three cities. This evidence, taken as a whole, points to a basic, underlying pattern-after two years, African Americans appear to benefit from vouchers, while others do not.

"Gains displayed by black children are most distinct during their first year in a private school; then the achievement advantage, relative to their peers in public schools, levels off." ***PACE*****, p. 8**

This claim is simply wrong. Our report shows quite clearly that while the impact of switching from a public to a private school for African Americans appeared to level off in New York after the first year, the impacts increased dramatically elsewhere-from -0.9 to 9.0 percentile points in D.C., and from 3.3 to 6.5 points in Dayton. When averaging across the three cities, the impacts for African Americans from year one to year two nearly doubled. For all other ethnic groups, significant impacts were not detected in any city in either year.

The demonstrated gains from using vouchers are limited to math, as reading performance appears to be "more difficult to budge." ***PACE*****, p. 8**

In no site after two years did we find significant differences between math and reading impacts. In New York, the year-two impacts for African Americans in math and reading were 4.1 and 4.5 national percentile points respectively; in Dayton, impacts were 5.3 and 7.6 points in math

and reading; and in D.C., they were 9.9 and 8.1 points. All of these impacts, except for math in Dayton, are statistically significant.

It is true that after the first year, significant impacts were limited to math in D.C.; only in New York did African Americans who switched from a public to a private school post significant and positive gains in both reading and math. When one examines the second year results, however, the differences between math and reading impacts disappear. Overall and in each of the three cities individually, test score impacts for African-American children in math and reading are comparable. Because of the similarities between math and reading scores, and in order to generate more stable estimates, our report focuses on the combined math and reading test scores.

Reported test score gains may be due to "the declining share of students who appeared for the standardized tests from years 1 and 2." *PACE*, p. 13

Not everyone in the test and control groups continued to participate in the evaluation two years later. This problem, which is encountered by virtually all evaluations of social interventions, is a valid concern. We did our best to locate and persuade as many families as possible to continue to participate in the evaluation, whether or not they had received a voucher. Still, for a variety of reasons, substantial numbers of students were not tested at the end of the second year.

We are reasonably confident, however, that this problem does not undermine the integrity of our findings. We obtained the test scores and background characteristics of virtually every student involved in the study at baseline, before they were randomly assigned to treatment or control groups. As our reports detail, these data reveal only minor differences between the second-year participants and non-participants in all three cities. To account for the modest differences we did observe, we weighted the data according to the predicted probability that each student, according to her baseline demographic characteristics, would attend the follow-up sessions.

In a randomized field trial, it is desirable to have similar response rates for test and control groups. If response rates differ noticeably, it is possible that the two groups participating in the study will no longer be comparable. Fortunately, in Dayton and D.C., the response rates for the test and control groups were essentially the same. Only in New York City did the students in the control group participate at a lower rate than the students offered a voucher-here the difference was 7 percentage points.

Because vouchers were randomly offered at baseline to test and control groups, results are unlikely to vary materially when one controls for family background characteristics. When assignment to the two groups is done at random, the two populations seldom differ significantly in their background characteristics. For this reason, our research team controlled only for baseline test scores in our original estimations. Since some critics of our evaluation have suggested that different results would be obtained if family background characteristics had been included as explicit controls in our models, we report for African Americans, in table 1 above,

results from estimations which control not only for initial test scores but also for mother's education, mother's employment status, family size, and whether or not the family received welfare. The estimated impacts on the test scores of African Americans of switching from a public to a private school in the three cities remain exactly the same—6.3 National Percentile Ranking (NPR) points, a statistically significant impact. Minor differences are observed when impacts within each individual city are estimated. When estimating effects in New York City without controlling for family background characteristics, the impact is estimated to be 4.4 NPR points; when family background controls are added, the impact is 4.2 NPR points. In Dayton, Ohio, when controls are introduced, the point estimate drops from 6.5 to 5.9 NPR points. And in Washington, D.C., the estimated impact increases from 9.0 to 9.1 NPR points. In two of the three cities, the estimated impacts, when controlled for family background characteristics, are statistically significant, and in the third, the impact just misses the standard threshold for statistical significance.

Table 1. Estimated Effects after Two Years of Switching from a Public to a Private School on African Americans' Combined Test Scores, With and Without Controls for Family Background Characteristics

	Private-School Impact, Original Results	Private-School Impact, Controlling for Family Background	p-value
Three-City Average Impact	6.3**	6.3**	[.012]
New York City	4.4*	4.2*	[.086]
Dayton, OH	6.5*	5.9	[.118]
Washington, D.C.	9.0***	9.1***	[.001]

* significant at .10 level, 2-tailed test; ** .05 level; *** .01 level. P-values reported in brackets. Weighted two-stage least squares regressions performed; treatment status used as instrument. All models control for baseline test scores, mother's education, employment status, whether or not the family receives welfare, and family size (missing case values for demographic variables estimated by imputation); NY model also includes lottery indicators. Impacts expressed in terms of national percentile rankings. Average three-city impact is based on effects observed in the three cities weighted by the inverse of the standard errors of the point estimates.

Need for Caution

As we emphasize in our report, which was presented and critiqued at the recent meeting of the American Political Science Association, one needs to exercise caution when drawing policy conclusions from our findings. These are only two-year results from fairly small, targeted pilot programs. Over the long run, results may become positive for all ethnic groups, or the observed effects of the program on African American students may dissipate altogether. And larger voucher programs may have quite different effects. Still, the weight of our evidence challenges

both voucher advocates and their critics: Positive impacts were consistently observed for African Americans, but not for anyone else.